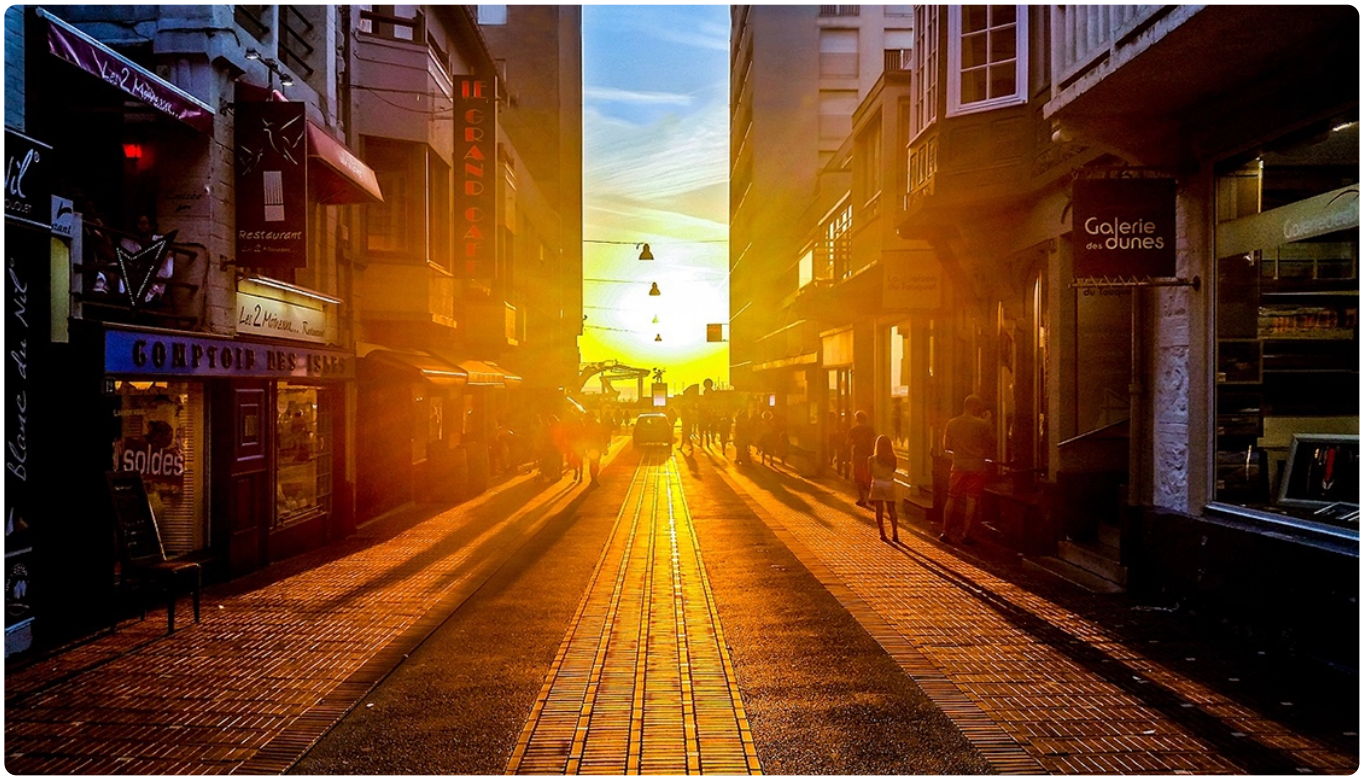
下载APP

12 | 精准定向：如何建立一个成熟的用户标签体系？

2022-03-25 郭谊

《大厂广告产品心法》 课程介绍 >



讲述：郭谊
时长 17:28 大小 16.00M

你好，我是郭谊。

在前两讲中，我已经带你了解了广告产品流量优化和物料生产的思路，做好这两点广告产品的点击率和转化率会有很大的提升。那除了流量和物料以外，你想一下，我们还可以从哪方面来优化广告效果呢？

没错，就是精准定向。

什么是精准定向广告？

> 精准定向广告：充分利用各种新式媒体，通过网络定向技术，以精准的渠道将营销信息传达给准确的目标受众群体。

通俗地说，**精准定向广告就是在对的地方，对的时间，向对的人，传递对的品牌和产品信息。**

精准定向广告有两大根本性的好处：

从购买广告位转向购买目标人群：无论购买了多少不同的广告位，只向目标人群展现广告主的广告，真正实现广告主想要与目标消费者进行有效沟通的梦想。

增加有效曝光，降低投放成本：所谓有效曝光，就是向广告主的目标人群所展示的广告曝光；减少了向非目标人群展示广告曝光的概率。

在广告产品的发展历程中，精准定向经历了以内容定向为主和以用户定向为主两个阶段。

> 内容定向：通过内容属性对广告进行定向投放，包括频道定向、上下文定向、搜索关键词定向等。

内容定向时代，做得最好的平台是 Google 谷歌。Google 的 Adwords（类似于百度搜索关键词广告）和 Adsense（类似于百度网盟广告）最初就是利用对网页内容的分析来实现精准定向。可以说，搜索关键词是内容定向里最精准的了。虽然内容定向广告的效果高于传统的非定向通投广告，但是在搜索关键词定向之外的领域，还无法满足广告主希望“抓住”目标人群的根本诉求，所以用户定向应运而生。

> 用户定向：通过用户属性、用户行为等用户标签体系对广告进行定向投放。

在国外，雅虎于 2007 年推出了根据用户在雅虎平台上的行为习惯及地理信息来进行投放的 SmartAds，可以说是用户定向广告的雏形。而 Facebook 利用它的社交产品的强账号体系收集了用户填写的性别、生日、职业等信息，在 2007 年晚些时候推出了用户定向广告。

目前，用户定向已经成为主流，并非所有的产品平台都是内容型的产品，但是所有的产品平台都要有流量才能做广告，而流量背后的实质就是用户。移动互联网和智能手机的广泛应用，又让用户定向有了更多的数据和技术基础。

质、量难两全：精准定向和海量曝光之间的矛盾和统一

互联网广告产品的头部玩家们，在展示自身的平台优势时，常提及的两个关键词就是：海量曝光和精准定向。不信我们可以打开百度广告、腾讯广告这些头部玩家的官方网站看一下，它们都强调了海量曝光和精准定向。

图1 百度和腾讯网站宣传页面展示

其实，海量曝光和精准定向之间存在着天然的矛盾。为什么这么说呢？

因为精准定向本身就要筛选出符合要求的人群，用户定位越精准，那曝光量相应的也就会越少。但是，海量曝光和精准定向之间又是统一的。因为只有拥有海量曝光的平台，才能满足精准定向广告的投放需求，所以精准定向是建立在海量曝光平台上的一项技术，同时我们可以根据自己的需求，张弛有度地使用精准定向投放广告。

精准定向实现的前提——建立用户标签体系

如果想要实现精准定向，我们首先要做的就是建立一个有效的用户标签体系。那什么是用户标签体系呢？

> 用户标签体系：基于大量的、多维度的用户数据，结合算法与建模，给每个用户打上人口学属性、行为属性、用户偏好等诸多类型的标签，多种标签可以进行组合和交叉，并

> 可以通过这些标签或者标签体系实现个性化内容推荐、精准定向广告、精准营销等目标。

讲到这里，想必你已经明白了，一个有效的用户标签体系的建立，是我们真正实现精准定向广告的前提。那么，我们应该如何建立一个有效的用户标签体系呢？

建立用户标签体系，可以分为以下几步：

1. 设计标签体系结构

目标先行，事半功倍。如果你要建立一个崭新的用户标签体系，可以从设计标签体系的结构开始。用户标签有不同的分类方法。这里给你介绍三种典型的分类方法：

第一种，按照标签属性来看，用户标签可以分为静态属性标签和动态属性标签：

> 静态属性标签：长期甚至永远有效的标签，例如用户的性别、出生日期等；
>
> 动态属性标签：需要定期更新才能保证有效性的标签，比如用户的活跃情况、购买情况等。

第二种，按照数据提取维度来看，用户标签可以分为事实标签、模型标签和预测标签：

> 事实标签：从原始数据中提取的标签，例如通过用户注册时填写的信息获取性别、生日、星座等。
>
> 模型标签：需要定义规则、建立模型计算得出的标签，例如用户支付意愿、推荐意愿等。
>
> 预测标签：参考已有事实数据，预测用户的行为或偏好，例如用户 A 的历史购物行为与用户群体 B 相似，使用协同过滤算法，可以预测用户 A 也喜欢群体 B 喜欢的某些商品。

我自己常用的是第三种方法，按照标签和用户的关系，把用户标签分为用户人口学属性、用户行为、用户偏好、其他等四大类：

用户人口学属性：例如性别、年龄、学历、职业、婚姻状况、居住城市等，这类标签是最基本的，也是最首要的；

用户行为：包括线上行为和线下行为。线上行为，例如用户是否浏览新闻资讯、刷抖音、打游戏、使用理财服务等。线下行为包括用户是否有车、是否经常购物等等；

用户偏好：用户对各类商品的购买兴趣，例如汽车、母婴、金融等等。

一般我们会设计两到三层的用户标签体系结构，一方面用来指导后续工作，另一方面这也是我们对用户标签最终形态的需求和体现。

在设计标签体系的时候，我们要注意两点：

一是标签最细粒度要能够直接支持精准广告相关业务，同时支持对应标签实例的规则自定义；

二是不同的标签要能够自由地组合为新的标签，同时支持标签间的关系、权重自定义。

2. 采集数据源

巧妇难为无米之炊。想要建立有效的用户标签体系，首先我们必须长期采集大量的、可靠的、多维度的数据源。用户标签体系可能的数据来源包括：

来自用户上网终端的设备数据、地理位置数据、天气数据等；

来自网页端、App 端的用户行为数据；

用户注册账号时自行填写的个人信息数据；

线上广告投放积累的用户与广告的曝光、点击、转化等数据；

线下渠道积累的用户 CRM（客户管理）数据等等。

百度、腾讯、阿里等企业，之所以能够成为移动互联网时代第一批广告巨头，从根本上来说，就是因为它们都具有丰厚的数据源基础。它们各有特色：百度胜在搜索关键词数据，因此对用户的购买意向判断最准确。腾讯胜在社交数据，因此对用户的人口学属性如性别、年龄和互动行为数据积累最深厚；阿里胜在电商数据，因此可以追踪到用户从登陆平台到完成商品购买的闭环数据。

其中，用户行为等动态数据源可以通过数据埋点和无埋点两种方法来采集。埋点更传统，你需要预先想好需要收集哪些数据，然后让技术团队去采集。无埋点实际上也可以叫做全埋点，顾名思义，就是采集所有可能收集到的用户数据，之后按照需求提取使用。这两者并无绝对的优劣之分，你可以和技术团队沟通，根据实际情况选择最适合你们的方案。

3. 完善数据仓库

数据源采集之后是不能直接使用的。我们需要根据精准定向广告的业务需求，对数据源进行选择和封装，在数据仓库中存储和培育数据。我们通过 ETL 把元数据统一格式后装载到不同的数据仓库里进行存储：有的数据库里存储的是用户人口学属性的数据；有的数据库存储的是某些类别的用户行为数据；有的数据库存储的是广告位相关数据。

> ETL：英文 Extract-Transform-Load 的缩写，用来描述将数据从来源端经过抽取（extract）、转换（transform）、加载（load）至目的端的过程。

数据仓库的工作是由技术团队和工程团队完成的，具体怎样去组建数据仓库需要广告、技术与工程团队之间密切地交流、沟通。

4. 建设用户标签体系应用层

数据仓库完善之后，我们如果有特定的数据分析需求，就可以从数据仓库中取数进行分析了。但是，单次的人工分析显然不能满足精准定向广告的需求，所以我们要在数据仓库的基础上，建设应用层。到了这个阶段，其实就和我们的第一步联系起来了。

以腾讯广告的用户标签体系为例，广告主或代理公司在广告投放前进行设置时，通过对人口学标签、设备类标签、兴趣类标签、垂直行业标签、自定义类标签、行为类标签六大类标签下的二级甚至三级标签的勾选与组合，实现目标人群定向。

图2 腾讯广告标签体系

常见问题

在建设用户标签体系的过程中，我们会不断面对新的问题，解决新的问题。我选择了其中三个高频的问题，我们来分析、解答一下。

问题一：为什么要先建立性别、地域、年龄这些人口学属性的基础标签？

一是因为这些基础标签下的用户量最大，可以很好地满足广告主海量曝光的需求，也符合媒体大量售卖广告产品的需要。

二是因为广告主如果选择的标签粒度太细，例如对很多用户行为标签进行圈选和交叉，那么取了这些标签的交集之后，目标人群规模可能会太小。这很有可能漏掉很多有潜力转化的用户，所以我们也不建议这样做，而是推荐广告主优先尝试按基础标签圈选的人群进行投放，也可以同时对规模较小的精准人群进行投放，然后对比两种方式的效果，或者先放量，后续再做人群优化。

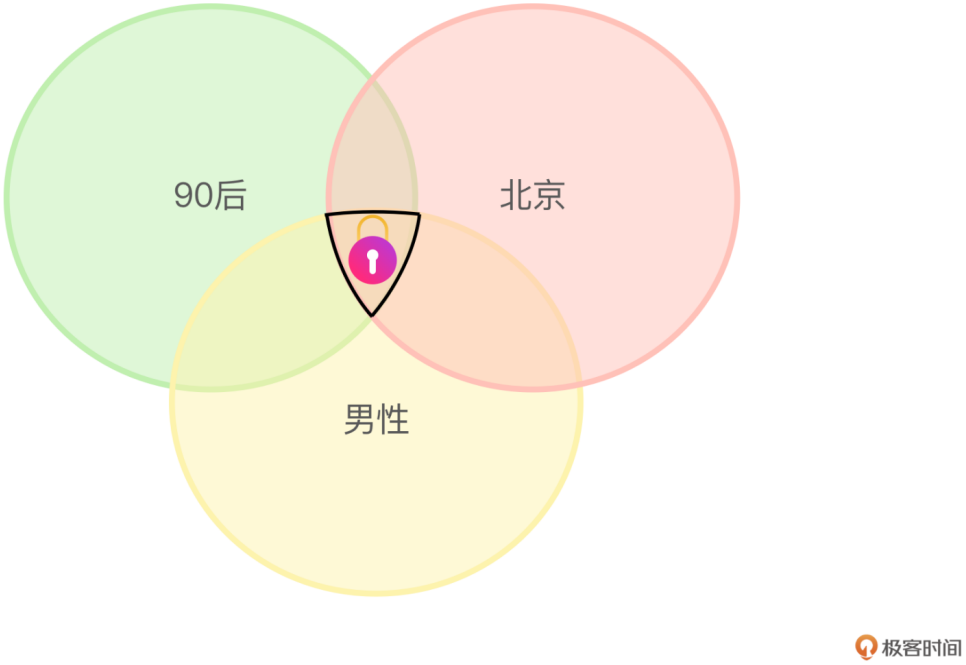


图3 标签交叉筛选示意图

三是因为性别、地域、年龄等标签，对各种类型的用户产品来说，都可以直接采集或者进行准确度较高的推测。

但是，根据我的经验，如果你去了一个初创的媒体广告平台，可能会发现它甚至连这些基础标签的清洗、封装和应用都没有做到位。在这种情况下，你首先要完善的就是这些在精准定向广告中应用最广泛的基础标签。

问题二：精准定向广告为什么要做用户兴趣标签？

用户兴趣标签，指的是基于用户的商业兴趣、泛娱乐兴趣或者语义兴趣等产生的标签。其中，商业兴趣标签最为重要。精准定向广告一开始只有用户人口学属性标签和用户行为标签，那后来为什么我们又要做这些兴趣标签呢？

原因是广告主，特别是效果广告主，一方面，总有性别、地域、年龄等人口学属性的标签满足不了的人群圈选需求；另一方面，就像前面说的，用户行为标签圈选出来的目标人群虽然可能更精准，但是往往规模较小，而且很多广告主对于目标人群的用户行为标签认知未必准确。

为了兼顾精准和投放量的需求，兴趣标签就顺势诞生了。我们在建设兴趣标签时，首先要考虑两点，一是广告主对这个兴趣标签的需求强度，基于这一点考虑，我们优先建设的人

群包往往是汽车、母婴、金融、电商、教育等广告主需求旺盛的行业；二是这个兴趣标签囊括的目标人群规模，如果兴趣标签下人群过少，那就对放量意义不大，可以舍弃。

兴趣标签的建设过程如下：

1. 首先，产品经理会根据业务反馈，把兴趣人群可能涉及的各个数据维度整理出来，例如搜索了关键词的人群、购买了商品的人群、浏览了商品信息的人群等等，作为行业知识输出给算法团队。
2. 算法团队参考这些维度进行数据挖掘与建模，把兴趣人群打成人群包。如果人群包的用户量级、精确率、召回率都符合精准定向广告投放的需求，我们就可以让销售联系广告主进行投放测试了。

> 精确率：该商品人群包中有多少用户是真正对商品感兴趣的人群。
> 召回率：对商品感兴趣的人群，有多大比例被包含在这个人群包里。

在兴趣标签建设完善后，我们应该大力推动广告主使用兴趣标签，这样不仅广告投放省时省事，而且也有利于广告产品的规模化售卖。

问题三：数据隐私权对用户标签体系是否会有影响？

近年来，国内和国际上对用户隐私的立法及相关保护都在推进中。

> 2021 年 8 月 20 日，第十三届全国人民代表大会常务委员会通过了《中华人民共和国个人信息保护法》。法律中明确规定， 处理个人信息应当具有明确、合理的目的，并应当与处理目的直接相关，采取对个人权益影响最小的方式；收集个人信息，应当限于实现处理目的的最小范围，不得过度收集个人信息等等。

这些立法必然会影响到各大媒体及广告平台对用户数据的采集、使用及告知方式。用户数据隐私是一个非常复杂的问题。作为广告产品的从业人员和关注者，我们务必要关注相关的法规政策的发展，还有业界最新的应对之策。

Lookalike：用户标签的拓展应用

最后，我们来聊一聊用户标签的拓展应用——Lookalike。

> Lookalike：字面意思是指一个人或事物和另外的人或事物非常相似。该技术又叫做“人群拓展”，指利用广告主第一方数据，基于少量的种子用户，通过大数据分析和机器学习拓展出和种子相似的用户人群。而这些拓展出的相似人群，同时也有很大可能会成为新的目标人群（比如，App 的下载激活、商品的收藏购买，目标粉丝的扩展等等）。

为了让你理解得更透彻，我给你举个例子。奔驰汽车给了腾讯广告一批它的线下 CRM 客户数据，主要是手机号。腾讯把这批手机号和自己平台上的用户标签进行匹配，发现奔驰汽车的这批人群在用户人口学属性、用户行为、社交关系链上体现出很多共同点。除了这批人群，腾讯平台上还有很多具备类似特征的相似人群，我们推测这些人群因为和奔驰已有的客户人群标签有相似性，因此对奔驰汽车感兴趣的可能性较高。所以就对这些相似人群投放了广告，取得了不错的点击率效果数据。

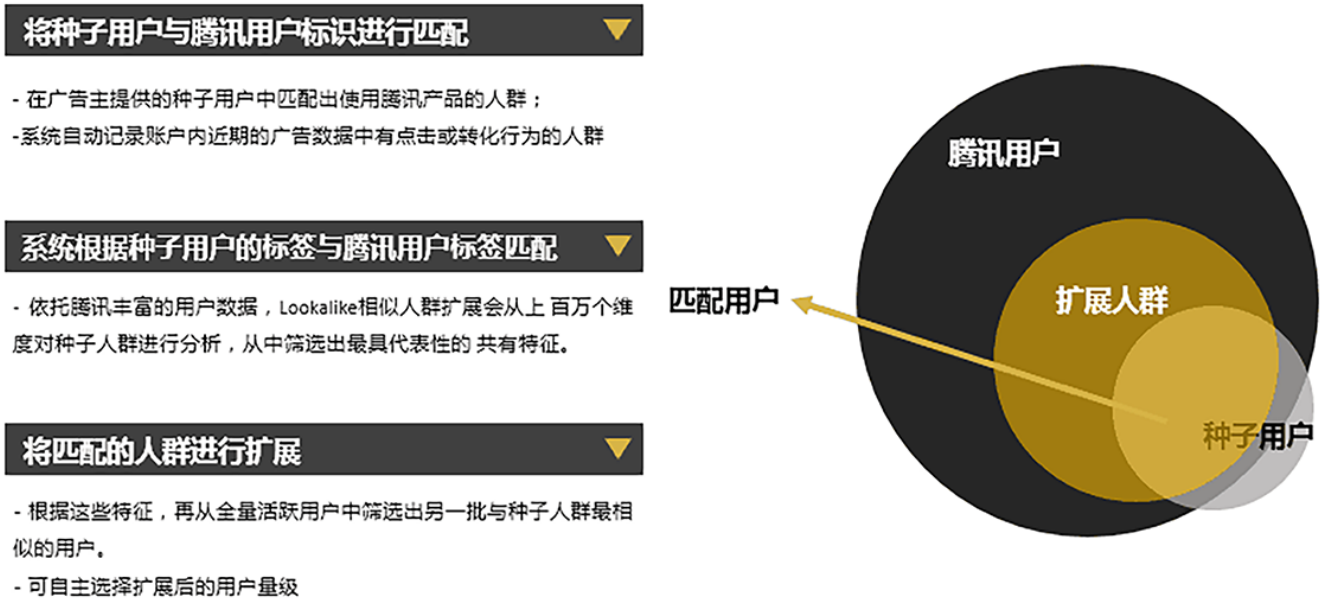
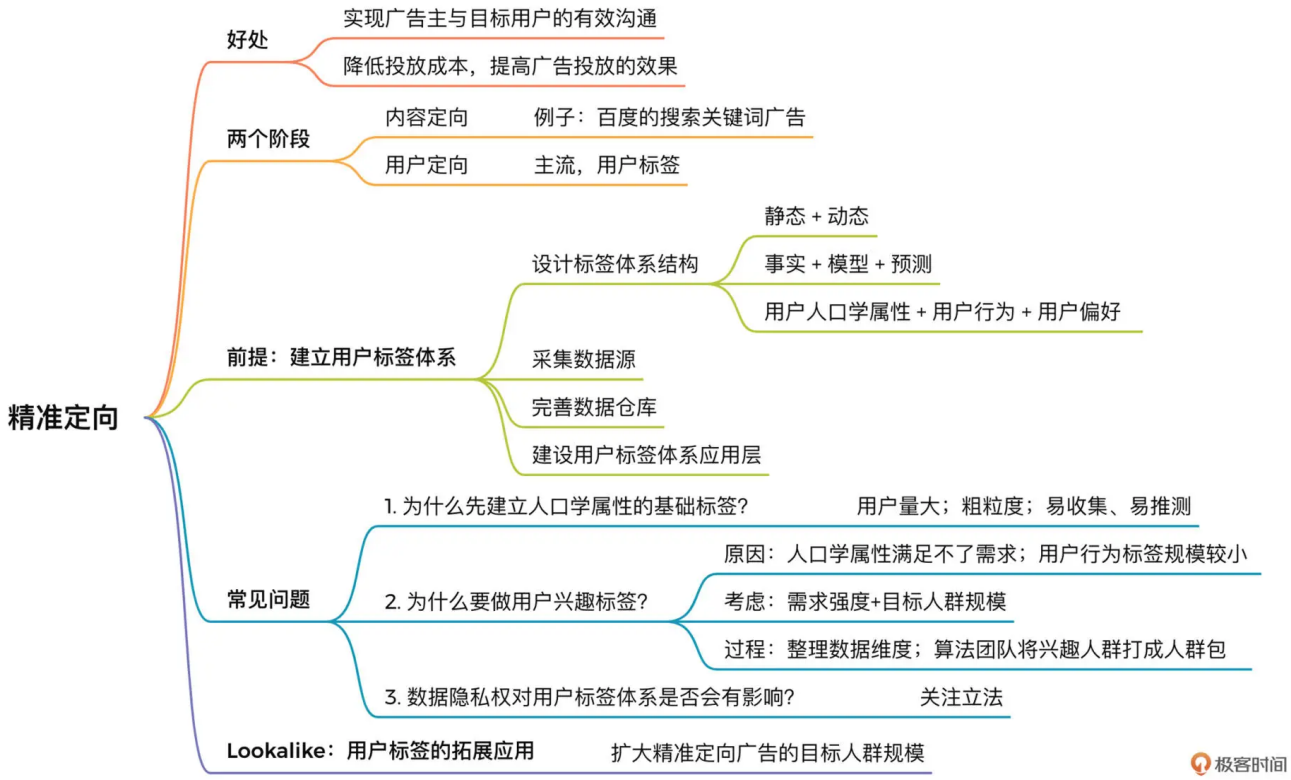


图4 腾讯使用lookalike技术扩展投放人群

Lookalike 经常用来扩大精准定向广告目标人群的规模，是一种非常有效的手段，能够帮助广告主挖掘出更多新鲜的潜在目标用户，最终提高广告投放的效果。

重点回顾



在这一讲中，我们从精准定向广告的定义出发，聚焦精准定向能力的核心来源——用户标签体系的相关知识。了解了建立用户标签体系的四部曲：设计标签体系结构、采集数据源、完善数据仓库和建设应用层。值得注意的是，我们一定要从业务需求出发设计合理的标签体系，做到有的放矢。

除此之外，我们在建设用户标签体系时，要优先建立人口学属性等基础标签，后续再进行精细化处理，建立用户兴趣标签。另外，我们还可以另辟蹊径，使用 Lookalike 人群拓展技术，它能够帮助我们扩大投放目标，触达新鲜人群。在整个过程中，因为涉及了用户隐私，所以要时刻关注数据隐私权方面的政策。

最后，我还需要提醒你：只有拥有海量曝光或者海量数据的媒体或者广告平台，才能真正做好精准定向广告产品，因为只有精准，没有规模的称不上真正的广告产品。同时，数据源仅仅是规模大也不行，还必须有丰富的维度。就像你手里有 10 万个手机号，没有任何标签信息，除了打违法的骚扰电话，还能怎么做广告呢？

实战演练

最后，我们来开动脑筋，处理两个实际问题吧！

某广告主为了推广自己的产品，在媒体 A 上投放广告，要求广告受众是生活在一线城市、年龄为 25~35 岁的女性，购买量为 1000 个 CPM。媒体 A 使用自身的用户标签和人群定

向系统进行广告投放。因为媒体 A 认为自己标签体系的准确度约为 92%。请问 A 为了实现针对目标人群的有效投放，实际投放了多少个 CPM?

该广告主自行向第三方数据监测服务公司购买了 15 万样本的数据。广告投放完毕之后，广告主发现该样本库中有 1 万个样本被媒体 A 投放了自己的广告，这其中有 8000 个样本确实是生活在一线城市、年龄为 25~35 岁的女性。请问根据该样本库结论，媒体 A 的人群定向准确性是多少？广告主要求按照第三方监测的结果给 A 付款。结合上一个问题中 A 实际投放的 CPM 数量，广告主愿意支付的 CPM 数量是多少个？

好了，今天的内容就到这里，欢迎你在评论区写出你的答案和我讨论，也欢迎你把这节课分享给身边的朋友，我们下节课再见！

分享给需要的人，Ta订阅超级会员，你最高得 50 元

Ta单独购买本课程，你将得 20 元

生成海报并分享

赞 2 提建议

© 版权归极客邦科技所有，未经许可不得传播售卖。 页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇 11 | 物料生产：如何满足广告主的创意需求？

精选留言 (3)

写留言

海平

2022-03-27

1.1087个CPM 2.80% 870个CPM

展开

viper

2022-03-25

1000/0.92=1086.956522 个cpm 收费 1086×0.8=868.8个cpm的钱

展开

作者回复: 谢谢viper同学的回答🌷咱们等等其他同学

F.J

2022-03-25

1087个CPM；80%；800CPM

作者回复: 谢谢F.J🌷咱们再看看其他同学的回答